

RECEIVED

Page 1 of 7

SEP 01 2000

TECH CENTER 1600

1649

#25

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/08/894,356CDATE: 08/24/2000  
TIME: 10:47:54Input Set : A:\1560-308.app  
Output Set: N:\CRF3\08232000\H894356C.raw

ENTERED

## SEQUENCE LISTING

4 (1) GENERAL INFORMATION:  
6 (i) APPLICANT: ASHIKARI, Toshihiko  
7 TANAKA, Yoshikazu  
8 FUJIWARA, Hiroyuki  
9 NAKAO, Masahiro  
10 FUKUI, Yuko  
11 SAKAKIBARA, Keiko  
12 MIZUTANI, Masako  
13 KUSUMI, Takaaki  
15 (ii) TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ACYL  
16 GROUP TRANSFER ACTIVITY  
18 (iii) NUMBER OF SEQUENCES: 31  
20 (iv) CORRESPONDENCE ADDRESS:  
21 (A) ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
22 (B) STREET: 1737 King Street, Suite 500  
23 (C) CITY: Alexandria  
24 (D) STATE: Virginia  
25 (E) COUNTRY: United States  
26 (F) ZIP: 22314-2756  
28 (v) COMPUTER READABLE FORM:  
29 (A) MEDIUM TYPE: Floppy disk  
30 (B) COMPUTER: IBM PC compatible  
31 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
32 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
34 (vi) CURRENT APPLICATION DATA:  
C--> 35 (A) APPLICATION NUMBER: US/08/894,356C  
C--> 36 (B) FILING DATE: 18-Aug-1997  
37 (C) CLASSIFICATION:  
51 (vii) PRIOR APPLICATION DATA:  
40 (A) APPLICATION NUMBER: JP 7-67159  
41 (B) FILING DATE: 17-FEB-1995  
44 (A) APPLICATION NUMBER: JP 7-196915  
45 (B) FILING DATE: 29-JUN-1995  
48 (A) APPLICATION NUMBER: JP 8-46534  
49 (B) FILING DATE: 30-JAN-1996  
52 (A) APPLICATION NUMBER: WO PCT/JP96/00348  
53 (B) FILING DATE: 16-FEB-1996  
55 (viii) ATTORNEY/AGENT INFORMATION:  
56 (A) NAME: Meuth, Donna M.  
57 (B) REGISTRATION NUMBER: 36,607  
58 (C) REFERENCE/DOCKET NUMBER: 001560-308  
60 (ix) TELECOMMUNICATION INFORMATION:  
61 (A) TELEPHONE: (703) 836-6620  
62 (B) TELEFAX: (703) 836-2021  
65 (2) INFORMATION FOR SEQ ID NO: 1:  
67 (i) SEQUENCE CHARACTERISTICS:

RECEIVED

RAW SEQUENCE LISTING                      DATE: 08/24/2000  
 PATENT APPLICATION: US/08/894,356C        TIME: 10:47:54

Input Set : A:\1560-308.app  
 Output Set: N:\CRF3\08232000\H894356C.raw

TECH GEN

```

68      (A) LENGTH: 1703 base pairs
69      (B) TYPE: nucleic acid
70      (C) STRANDEDNESS: double
71      (D) TOPOLOGY: linear
73      (ii) MOLECULE TYPE: cDNA to mRNA
75      (iii) HYPOTHETICAL: NO
77      (iv) ANTI-SENSE: NO
79      (vi) ORIGINAL SOURCE:
80          (A) ORGANISM: Gentiana triflora var. japonica
81          (F) TISSUE TYPE: petal
83      (vii) IMMEDIATE SOURCE:
84          (A) LIBRARY: cDNA library
85          (B) CLONE: pGAT4
87      (ix) FEATURE:
88          (A) NAME/KEY: CDS
89          (B) LOCATION: 6..1412
92      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
94 TCATT ATG GAG CAA ATC CAA ATG GTG AAG GTT CTT GAA AAA TGC CAA      47
95      Met Glu Gln Ile Gln Met Val Lys Val Leu Glu Lys Cys Gln
96          1          5          10
98 GTT ACA CCA CCA TCT GAC ACA ACA GAT GTC GAG TTA TCG CTA CCG GTA      95
99 Val Thr Pro Pro Ser Asp Thr Thr Asp Val Glu Leu Ser Leu Pro Val
100 15          20          25          30
102 ACA TTC TTC GAT ATC CCC TGG TTG CAC TTG AAT AAG ATG CAG TCC CTT      143
103 Thr Phe Phe Asp Ile Pro Trp Leu His Leu Asn Lys Met Gln Ser Leu
104          35          40          45
106 CTG TTT TAC GAC TTT CCG TAC CCA AGA ACA CAT TTC TTG GAC ACT GTT      191
107 Leu Phe Tyr Asp Phe Pro Tyr Pro Arg Thr His Phe Leu Asp Thr Val
108          50          55          60
110 ATC CCT AAT CTT AAG GCC TCT TTG TCT CTC ACT CTA AAA CAC TAC GTT      239
111 Ile Pro Asn Leu Lys Ala Ser Leu Ser Leu Thr Leu Lys His Tyr Val
112          65          70          75
114 CCG CTT AGC GGA AAT TTG TTG ATG CCG ATC AAA TCG GGC GAA ATG CCG      287
115 Pro Leu Ser Gly Asn Leu Leu Met Pro Ile Lys Ser Gly Glu Met Pro
116          80          85          90
118 AAG TTT CAG TAC TCC CGT GAT GAG GGC GAC TCG ATA ACT TTG ATC GTT      335
119 Lys Phe Gln Tyr Ser Arg Asp Glu Gly Asp Ser Ile Thr Leu Ile Val
120 95          100          105          110
122 GCG GAG TCT GAC CAG GAT TTT GAC TAC CTT AAA GGT CAT CAA CTG GTA      383
123 Ala Glu Ser Asp Gln Asp Phe Asp Tyr Leu Lys Gly His Gln Leu Val
124          115          120          125
126 GAT TCC AAT GAT TTG CAT GGC CTT TTT TAT GTT ATG CCA CGG GTT ATA      431
127 Asp Ser Asn Asp Leu His Gly Leu Phe Tyr Val Met Pro Arg Val Ile
128          130          135          140
130 AGG ACC ATG CAA GAC TAT AAA GTG ATC CCG CTC GTA GCC GTG CAA GTA      479
131 Arg Thr Met Gln Asp Tyr Lys Val Ile Pro Leu Val Ala Val Gln Val
132          145          150          155
134 ACC GTT TTT CCT AAC CGT GGC ATA GCC GTG GCT CTG ACG GCA CAT CAT      527
135 Thr Val Phe Pro Asn Arg Gly Ile Ala Val Ala Leu Thr Ala His His

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/08/894,356C  
 DATE: 08/24/2000  
 TIME: 10:47:54

Input Set : A:\1560-308.app  
 Output Set: N:\CRF3\08232000\H894356C.raw

136	160	165	170	
138 TCA ATT GCA GAT GCT AAA AGT TTT GTA ATG TTC ATC AAT GCT TGG GCC	575			
139 Ser Ile Ala Asp Ala Lys Ser Phe Val Met Phe Ile Asn Ala Trp Ala				
140 175 180 185 190				
142 TAT ATT AAC AAA TTT GGG AAA GAC GCG GAC TTG TTG TCC GCG AAT CTT	623			
143 Tyr Ile Asn Lys Phe Gly Lys Asp Ala Asp Leu Leu Ser Ala Asn Leu				
144 195 200 205				
146 CTT CCA TCT TTC GAT AGA TCG ATA ATC AAA GAT CTG TAT GGC CTA GAG	671			
147 Leu Pro Ser Phe Asp Arg Ser Ile Ile Lys Asp Leu Tyr Gly Leu Glu				
148 210 215 220				
150 GAA ACA TTT TGG AAC GAA ATG CAA GAT GTT CTT GAA ATG TTC TCT AGA	719			
151 Glu Thr Phe Trp Asn Glu Met Gln Asp Val Leu Glu Met Phe Ser Arg				
152 225 230 235				
154 TTT GGA AGC AAA CCC CCT CGA TTC AAC AAG GTA CGA GCT ACA TAT GTC	767			
155 Phe Gly Ser Lys Pro Pro Arg Phe Asn Lys Val Arg Ala Thr Tyr Val				
156 240 245 250				
158 CTC TCC CTT GCT GAA ATC CAG AAG CTA AAG AAC GTA CTG AAT CTC	815			
159 Leu Ser Leu Ala Glu Ile Gln Lys Leu Lys Asn Lys Val Leu Asn Leu				
160 255 260 265 270				
162 AGA GGA TCC GAA CCG ACA ATA CGT GTA ACG ACG TTC ACA ATG ACG TGT	863			
163 Arg Gly Ser Glu Pro Thr Ile Arg Val Thr Thr Phe Thr Met Thr Cys				
164 275 280 285				
166 GGA TAC GTA TGG ACA TGC ATG GTC AAA TCA AAA GAT GAC GTC GTA TCA	911			
167 Gly Tyr Val Trp Thr Cys Met Val Lys Ser Lys Asp Asp Val Val Ser				
168 290 295 300				
170 GAG GAA TCA TCG AAC GAC GAA AAT GAG CTC GAG TAC TTC AGT TTT ACA	959			
171 Glu Glu Ser Ser Ser Asn Asp Glu Asn Glu Leu Glu Tyr Phe Ser Phe Thr				
172 305 310 315				
174 GCG GAT TGC CGA GGA CTT CTG ACG CCC CCG TGT CCG CCT AAC TAC TTT	1007			
175 Ala Asp Cys Arg Gly Leu Leu Thr Pro Pro Cys Pro Pro Asn Tyr Phe				
176 320 325 330				
178 GGC AAC TGT CTT GCG TCA TGC GTT GCA AAA GCA ACA CAT AAA GAG TTA	1055			
179 Gly Asn Cys Leu Ala Ser Cys Val Ala Lys Ala Thr His Lys Glu Leu				
180 335 340 345 350				
182 GTT GGG GAT AAA GGG CTT CTT GTT GCA GTT GCA GCT ATT GGA GAA GCC	1103			
183 Val Gly Asp Lys Gly Leu Leu Val Ala Val Ala Ala Ile Gly Glu Ala				
184 355 360 365				
186 ATT GAA AAG AGG TTG CAC AAC GAA AAA GGC GTT CTT GCA GAT GCA AAA	1151			
187 Ile Glu Lys Arg Leu His Asn Glu Lys Gly Val Leu Ala Asp Ala Lys				
188 370 375 380				
190 ACT TGG TTA TCG GAA TCT AAT GGA ATC CCT TCA AAA AGA TTT CTC GGG	1199			
191 Thr Trp Leu Ser Glu Ser Asn Gly Ile Pro Ser Lys Arg Phe Leu Gly				
192 385 390 395				
194 ATT ACC GGA TCG CCT AAG TTC GAT TCG TAT GGT GTA GAT TTT GGA TGG	1247			
195 Ile Thr Gly Ser Pro Lys Phe Asp Ser Tyr Gly Val Asp Phe Gly Trp				
196 400 405 410				
198 GGA AAG CCT GCA AAA TTT GAC ATT ACC TCT GTT GAT TAT GCA GAA TTG	1295			
199 Gly Lys Pro Ala Lys Phe Asp Ile Thr Ser Val Asp Tyr Ala Glu Leu				
200 415 420 425 430				

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/08/894,356C

DATE: 08/24/2000  
 TIME: 10:47:54

Input Set : A:\1560-308.app  
 Output Set: N:\CRF3\08232000\H894356C.raw

```

202 ATT TAT GTG ATT CAG TCC AGG GAT TTT GAA AAA GGT GTG GAG ATT GGA      1343
203 Ile Tyr Val Ile Gln Ser Arg Asp Phe Glu Lys Gly Val Glu Ile Gly
204                               435                               440       445
206 GTA TCA TTG CCT AAG ATT CAT ATG GAT GCA TTT GCA AAA ATC TTT GAA      1391
207 Val Ser Leu Pro Lys Ile His Met Asp Ala Phe Ala Lys Ile Phe Glu
208                               450                               455       460
210 GAA GGC TTT TGC TCT TTG TCA TAGTCTCTTT AATAGAACCA TATTTGCTGC      1442
211 Glu Gly Phe Cys Ser Leu Ser
212                               465
214 AATAAAGTAC CAAGTCCTTT AGTAACACTA CACCAAACCC TACTTTCGAG GCGGGAACAC      1502
216 CACAACGAGG TTCAATCACT AGAAGGTTGT ACTTCATAAA TTCCAGAGGT CGAATATACA      1562
218 CCGTTGTCCT CTGAAAAGTT GAACCTCACA CCTGACATGG TGTTACGATA GGTATTGTAT      1622
220 AATGCCATTA TATACTTCCA TAAAGTATCC TATGCAATAG AGAACATGTT ATGTGTAA      1682
222 AAAAAAAAAA AAAAAAAAAA A
225 (2) INFORMATION FOR SEQ ID NO: 2:
227 (i) SEQUENCE CHARACTERISTICS:
228 (A) LENGTH: 1622 base pairs
229 (B) TYPE: nucleic acid
230 (C) STRANDEDNESS: double
231 (D) TOPOLOGY: linear
233 (ii) MOLECULE TYPE: cDNA to mRNA
235 (iii) HYPOTHETICAL: NO
237 (iv) ANTI-SENSE: NO
239 (vi) ORIGINAL SOURCE:
240 (A) ORGANISM: Gentiana triflora va. japonica
241 (F) TISSUE TYPE: petal
243 (vii) IMMEDIATE SOURCE:
244 (A) LIBRARY: cDNA library
245 (B) CLONE: pGAT106
247 (ix) FEATURE:
248 (A) NAME/KEY: CDS
249 (B) LOCATION: 35..1471
252 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
254 GAACCATTGA ATCCAATTAA TCTGATTAT TAAG ATG GCA GGA AAT TCC GAG      52
255                               1                               5
256                               Met Ala Gly Asn Ser Glu
258 GAT ATC AAA GTT CTT GAG AAA TGC CGT GTT GCG CCA CCA CCG GAC GCC      100
259 Asp Ile Lys Val Leu Glu Lys Cys Arg Val Ala Pro Pro Pro Asp Ala
260                               10                               15       20
262 GTC GCC GAG TTT ACA GTC CCA CTG TCG TTT TTC GAC ATG CGA TGG TTG      148
263 Val Ala Glu Phe Thr Val Pro Leu Ser Phe Phe Asp Met Arg Trp Leu
264                               25                               30       35
266 ATC TCT GAT GCA GAA CAC CAT CTG CAT TTC TAC AGA TTC CGC CAT CCT      196
267 Ile Ser Asp Ala Glu His His Leu His Phe Tyr Arg Phe Arg His Pro
268                               40                               45       50
270 TGT CCC AAC TCT AAA TTT ATC ATT TCA TCC ATT AAA TCG TCC CTT TCC      244
271 Cys Pro Asn Ser Lys Phe Ile Ile Ser Ser Ile Lys Ser Ser Leu Ser
272                               55                               60       65       70
274 CTT GTT CTC AAA CAC TTT CTT CCG TTA GCC GGG AAT TTG ATT TGG CCG      292

```

RAW SEQUENCE LISTING                      DATE: 08/24/2000  
 PATENT APPLICATION: US/08/894,356C        TIME: 10:47:54

Input Set : A:\1560-308.app  
 Output Set: N:\CRF3\08232000\H894356C.raw

275	Leu	Val	Leu	Lys	His	Phe	Leu	Pro	Leu	Ala	Gly	Asn	Leu	Ile	Trp	Pro	
276				75						80				85			
278	GTA	GAT	TCC	TCC	GAT	AGA	ATG	CCG	GAG	TTG	CGT	TAC	AAG	AAA	GGG	GAC	340
279	Val	Asp	Ser	Ser	Asp	Arg	Met	Pro	Glu	Leu	Arg	Tyr	Lys	Lys	Gly	Asp	
280				90					95					100			
282	TCC	GTT	TCT	TTA	ACA	ATT	GCA	GAA	TCG	AGC	ATG	GAT	TTT	GAT	TAT	CTC	388
283	Ser	Val	Ser	Leu	Thr	Ile	Ala	Glu	Ser	Ser	Met	Asp	Phe	Asp	Tyr	Leu	
284				105					110					115			
286	GCC	GGA	GAT	CAT	CAG	AGG	GAT	TCT	TAT	AAA	TTC	AAC	GAT	TTG	ATT	CCG	436
287	Ala	Gly	Asp	His	Gln	Arg	Asp	Ser	Tyr	Lys	Phe	Asn	Asp	Leu	Ile	Pro	
288				120					125					130			
290	CAG	CTG	CCA	GAA	CCG	ATT	GTA	ACC	TCC	GGC	GAC	GAA	GTA	TTA	CCA	CTT	484
291	Gln	Leu	Pro	Glu	Pro	Ile	Val	Thr	Ser	Gly	Asp	Glu	Val	Leu	Pro	Leu	
292	135				140					145						150	
294	TTT	GCT	TTA	CAG	GTG	ACG	GTG	TTC	TCC	AAC	ACC	GGT	ATA	TGC	ATT	GGA	532
295	Phe	Ala	Leu	Gln	Val	Thr	Val	Phe	Ser	Asn	Thr	Gly	Ile	Cys	Ile	Gly	
296				155						160						165	
298	CGC	AAT	CTT	CAT	CAA	GTT	CTT	GGT	GAT	GCC	AGT	TCT	TTT	CTG	CAT	TTT	580
299	Arg	Asn	Leu	His	Gln	Val	Leu	Gly	Asp	Ala	Ser	Ser	Phe	Leu	His	Phe	
300				170						175				180			
302	AAT	AAA	TTA	TGG	GTT	TTG	GTT	GAC	AAA	TCC	AAT	GGA	GAT	TCA	TTA	AAG	628
303	Asn	Lys	Leu	Trp	Val	Leu	Val	Asp	Lys	Ser	Asn	Gly	Asp	Ser	Leu	Lys	
304				185					190					195			
306	TTC	CTT	CCA	CTT	TCT	TCT	CTA	CCT	ATG	TAC	GAC	AGA	TCT	GTG	GTG	CAA	676
307	Phe	Leu	Pro	Leu	Ser	Ser	Leu	Pro	Met	Tyr	Asp	Arg	Ser	Val	Val	Gln	
308				200					205					210			
310	GAT	CCA	TTT	CAT	ATT	CGT	CGA	AAA	ATC	TAC	AAT	GAA	AGA	AAA	CTG	CTC	724
311	Asp	Pro	Phe	His	Ile	Arg	Arg	Lys	Ile	Tyr	Asn	Glu	Arg	Lys	Leu	Leu	
312	215					220					225					230	
314	AAA	TCT	CAG	GGC	ACA	CCT	ACT	GTT	CTA	AAT	CCA	GCA	ATT	TCT	AAA	GAT	772
315	Lys	Ser	Gln	Gly	Thr	Pro	Thr	Val	Leu	Asn	Pro	Ala	Ile	Ser	Lys	Asp	
316				235						240						245	
318	GAA	GTT	CGA	GCC	ACC	TTC	ATC	CTA	CAC	CCT	ATT	GAT	ATC	ATG	AAG	CTC	820
319	Glu	Val	Arg	Ala	Thr	Phe	Ile	Leu	His	Pro	Ile	Asp	Ile	Met	Lys	Leu	
320				250						255						260	
322	AAG	AAA	TTC	ATT	TCG	TCA	AAA	AAT	CGC	AAC	TTA	ACC	GGT	AGT	AGT	AAT	868
323	Lys	Lys	Phe	Ile	Ser	Ser	Lys	Asn	Arg	Asn	Leu	Thr	Gly	Ser	Ser	Asn	
324				265					270					275			
326	TAT	AAT	CTG	TCA	ACT	TTC	ACG	GTG	ACA	TCT	GCA	CTG	ATC	TGG	ACA	TGC	916
327	Tyr	Asn	Leu	Ser	Thr	Phe	Thr	Val	Thr	Ser	Ala	Leu	Ile	Trp	Thr	Cys	
328				280					285					290			
330	TTG	TCG	AAA	TCA	TTA	GAC	ACC	GTC	GTA	AGA	GAG	AAG	GTG	GAA	GAG	GAT	964
331	Leu	Ser	Lys	Ser	Leu	Asp	Thr	Val	Val	Arg	Glu	Lys	Val	Glu	Glu	Asp	
332	295					300					305					310	
334	AAA	CAT	GCA	GCA	AAC	TTA	TGT	GCT	TTC	ATC	AAC	TGC	CGA	CAA	CGT	TTT	1012
335	Lys	His	Ala	Ala	Asn	Leu	Cys	Ala	Phe	Ile	Asn	Cys	Arg	Gln	Arg	Phe	
336				315						320						325	
338	GCT	CCG	CCG	ATA	CCT	CAA	AAT	TAC	TTT	GGA	AAT	TGC	ATA	GTG	CCT	TGT	1060
339	Ala	Pro	Pro	Ile	Pro	Gln	Asn	Tyr	Phe	Gly	Asn	Cys	Ile	Val	Pro	Cys	

## VERIFICATION SUMMARY

DATE: 08/24/2000

PATENT APPLICATION: US/08/894,356C

TIME: 10:47:55

Input Set : A:\1560-308.app

Output Set: N:\CRF3\08232000\H894356C.raw

L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:867 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:876 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:880 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:884 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:888 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:892 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:896 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:900 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:904 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:908 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:912 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:916 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:920 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:924 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:928 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:932 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:936 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:940 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:944 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:948 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:952 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:956 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:960 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:964 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:968 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:972 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:976 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:980 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:1402 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:29  
L:1442 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:31